

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA

RICHARD KADREY, et al.,
Plaintiffs,
v.
META PLATFORMS, INC.,
Defendant.

Case No. 23-cv-03417-VC

QUESTIONS FOR HEARING

The parties should consider the following non-exclusive list of questions in advance of tomorrow's hearing:

1. It's difficult, from reading Meta's brief, to identify the "secondary use" that should be compared to the typical use of a book to assess whether the secondary use is transformative. At times, Meta appears to be contending that its use of the books is transformative because of what Llama is ultimately capable of producing. See Meta's 1st Br. at 14–15 ("It would be inconsistent with the purpose of copyright to allow the limited monopoly conferred by Plaintiffs' copyrights to interfere with the development of a new technology as innovative and quintessentially transformative as Llama.") Perhaps that's consistent with the Supreme Court's analysis in *Google v. Oracle*. But if that were the right way to examine whether a use is transformative in this context, then wouldn't it be fair use for a professor to download a pirated book, copy it, and give it to a brilliant student (rather than buying it for the student or requiring the student to buy it), knowing that the student will absorb the book (along with many other books) and use the knowledge to do something transformative? In the context of this case, is it better to think of the "secondary use" of the book as simply training the language model with it, without regard to

what the language model will ultimately create? In other words, is the real question whether using the book to train the language model is sufficiently different from a human reading the book so as to make it transformative, without regard to what the language model ultimately enables people to produce after the copying?

2. Assume that instead (or in addition) the Court should be looking to what the language model enables people to produce to assess whether the use is transformative. On this view, Meta contends that its use is transformative because the information is “used to enable Llama to perform functions and create outputs completely unrelated to, and different from,” the plaintiffs’ books. Meta’s 1st Br. at 17. But presumably Llama can also perform functions that are very related to, and very similar to, the functions performed by a plaintiff’s book. For example, presumably Llama can—drawing from the work of Rachel Louise Snyder and every other work it has ingested on the issue of domestic violence—write a detailed, *New Yorker*-style article on the roots of domestic violence and what must be done to combat it. If the copying enables the model to produce both comparable and non-comparable works, how does that affect the analysis?

3. This issue also goes to the fourth fair use factor. If Llama is able to write the above-referenced essay, isn’t that a substitute in the market for Snyder’s book? Or imagine that a company, without permission, feeds its language model every copyright-protected poem ever written, which gives the language model the ability to write millions of original poems. Or imagine feeding the language model every copyright protected magazine article ever written on the history of music, which enables the model to write innumerable articles on the history of music from all angles? In a situation like this, hasn’t the use of the copyright-protected poems or articles to train the model made created a “substitute in a market the copyright holder reasonably expected?” Meta’s 1st Br. at 24. Or, at least, hasn’t the use of the poems or articles created a serious likelihood that a substitute in the market will soon develop?

4. What does the summary judgment record say about the ability of Llama to do the things described above?

5. If the record allows an inference that Llama can do the things described above, does the record say anything about how doing those things has affected or would affect the market for the copied works?

6. Should works of fiction and non-fiction be thought of differently with respect to the fourth factor, in that it may be more likely for a language model to produce outputs that interfere with the market for original non-fiction works? Does the record say anything about this?

7. Are there cases where the secondary use was found to be transformative but the fair use doctrine was found not to apply (perhaps because of the fourth factor)?

8. There must be some difference, from a fair use standpoint, between downloading pirated works and using them to train AI versus lawfully acquiring the works and then using them to train AI. The plaintiffs' argument that this is dispositive in their favor seems wrong, but Meta seems equally wrong to argue that it's entirely irrelevant. Assuming the Court rejects both sides' arguments on this point, how should the issue be treated?

9. If companies are allowed to download pirated works to train their AI models, will that facilitate broader use of shadow libraries like LibGen? Does anything in the record speak to this question? How does the answer affect whether Meta's downloading (and not just its alleged uploading) of copyrighted works is fair use?

10. In terms of promoting creativity through fair use, is it typically a person using Llama to do the creating? Or is it Llama itself that's typically doing the creating? What does the record say about how Llama is commonly used in the real world? And does the distinction matter from a fair use standpoint?


11. What does the record say about the extent to which using copyright-protected material enhances Llama's creativity or effectiveness? Is this relevant to the fair use analysis?

12. Imagine a case where the evidence showed that: (1) allowing use of protected books to train an AI model would, to a degree, diminish the market for the copied works; and (2) disallowing use of protected books to train the AI model would, to a degree, diminish the effectiveness of the AI model's ability to generate high-quality output. In that scenario, would

the question of how easy or difficult it would be to obtain licenses to use the protected works for AI training be relevant to the fair use analysis?

IT IS SO ORDERED.

Dated: April 30, 2025



VINCE CHHABRIA
United States District Judge